Analysis of Protein Complexes from *Rhodopseudomonas palustris* by Mass Spectrometry

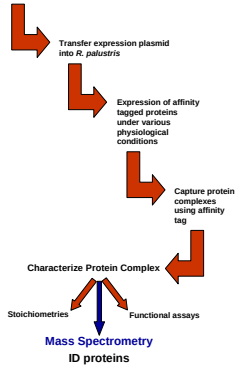
Gregory B. Hurst, Dale A. Pelletier, Robert L. Hettich, Keiji G. Asano, Michael B. Strader, David L. Tabb, Nathan C. VerBerkmoes, Trish K. Lankford, Linda J. Foote, Yisong Wang, Stephen J. Kennel, Frank W. Larimer
Oak Ridge National Laboratory, Oak Ridge, TN

OVERVIEW

- **Target proteins in *Rhodopseudomonas palustris* have been expressed as fusions with affinity labels to enable isolation of protein complexes.**
- **Component proteins of several complexes have been affinity-isolated and characterized by MS.**

INTRODUCTION

- The bacterial species *Rhodopseudomonas palustris* [1]
  - occurs widely in the environment
  - survives in a variety of conditions
    - light / dark
    - aerobic / anaerobic
- This species thus has the potential to express markedly different complements of proteins and protein complexes under different growth conditions.
- As part of a center funded by the U.S. Department of Energy Genomes To Life program [2,3], we are analyzing protein complexes from *R. palustris* by expressing target proteins as fusions with affinity tags to allow subsequent isolation of other proteins associated with the target [4].
- The MS analysis approach includes
  - Bottom-up analysis of affinity-isolated complexes (this poster)
  - Bottom-up proteomics of *R. palustris* proteins (poster ThPT 382, VerBerkmoes *et al.*)
  - Top-down analysis of intact proteins by FT-ICR-MS (poster MPU 390, Connelly *et al.*)
  - Integrated top-down and bottom-up characterization [5] of a conventionally-isolated complex (70S ribosome from *R. palustris*). [6] (poster ThPN 270, Strader *et al.*)



EXPERIMENTAL

- Selected *R. palustris* genes were cloned with affinity tags [7] and expressed in both *E. coli* and *R. palustris* using modified pDEST vectors (Invitrogen).
  - Tag contains both His6 and V5 epitope
  - C-terminal position
- Isolation of fusion proteins
  - Affinity purification with Ni-NTA agarose beads, followed by anti-V5 antibody agarose beads
  - Expression confirmed using 1-D PAGE and western blots.
- "Shotgun" analysis: analysis by mass spectrometry without prior gel separation [8]
  - trypsin digestion
  - Reverse-phase HPLC separation online with electrospray/quadrupole ion trap MS-MS
  - Protein ID's: Sequest [9] analysis of tandem mass spectral data using *R. palustris* database [1] .

RESULTS

- Plasmids encoding 42 affinity-tagged fusion proteins have been inserted into *R. palustris* (**Table 1**).
  - Choice of target proteins guided by proteomics measurements (poster ThPT 382, VerBerkmoes *et al.*)
- Multiple protein components of several complexes have been identified:
  - Iron nitrogenase (**Table 2**)
  - RNA polymerase (**Table 3**)
  - ATP synthase
  - GroESL
- Over 400 "background" (i.e., unlabeled) proteins were commonly observed in numerous LC-MS-MS experiments. **Figure 1** shows the most frequently detected of these.
- Based on this expanding library of background measurements, we are currently developing criteria for determining whether detection of a given protein in a particular isolation/mass spectrometry experiment represents a valid "hit."
  - One rough indication of a valid "hit" may be whether selected parameters for a detection (e.g. sequence coverage, number of unique peptides, total number of peptide spectra) significantly exceed those observed in the background for that protein.
  - As more measurements are made, the background library expands, and "hits" can be continually re-evaluated and refined.

Figure 1. Background proteins observed by LC-MS-MS in Affinity Isolation Experiments

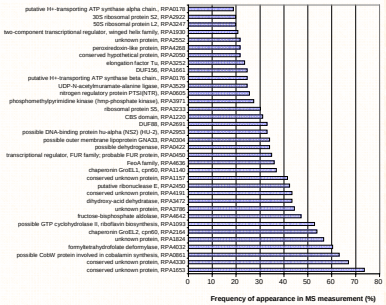


Table 1. Affinity-tagged proteins

Gene #	Name	Description	Typical Results: Unique peptides	Typical Results: % sequence coverage	# trials
RPA0176	atpD	ATP synthase beta subunit	35	79	7
RPA0190	sucD	succinyl-CoA synthetase alpha-subunit	18	60	1
RPA0191	sucC	succinyl-coA synthetase beta chain	16	60	1
RPA0192	mdh	malate dehydrogenase	41	76	1
RPA0245	ffh/fstY	signal recognition particle	1	1.4	1
RPA0962	hupS	uptake hydrogenase small subunit	1	5	2
RPA0963	hupL	uptake hydrogenase large subunit	45	67	6
RPA1141	groES 1	small subunit of GroESL molecular chaperone	9	65	3
RPA1548	puhA	H subunit of photosynthetic reaction center	ND	ND	2
RPA1980	cbbZ	phosphoglycolate phosphatase	ND	ND	1
RPA1981	cbbI, rpiA, rpi	ribose 5-phosphate isomerase	28	82	2
RPA2164	groEL 2	large subunit of GroESL molecular chaperone	62	72	9
RPA2165	groES 2	small subunit of GroESL molecular chaperone	6	41	4
RPA2336	RPA2336	unknown protein	26	73	3
RPA2339		possible iron response transcription regulator	ND	ND	1
RPA2405	dsrT	NAD+ ADP-ribosyltransferase	ND	ND	1
RPA2406	dsrD	dinitrogenase reductase activating	ND	ND	1
RPA2462	RPA2462	conserved unknown protein	2	23	1
RPA2867		pyruvate dehydrogenase E1 alpha subunit	ND	ND	1
RPA2967	glnA	glutamine synthetase I	36	56	1
RPA2969	RPA2969	unknown protein	8	40	2
RPA3147	clpA	endopeptidase Clp: ATP-binding chain A	ND	ND	1
RPA3226	rpoA	DNA-directed RNA polymerase alpha subunit	37	73	2
RPA3247	rplB	50S ribosomal protein L2	9	42	1
RPA3248	rplW	50S ribosomal protein L23	13	72	1
RPA3252	tuA, EF-Tu	elongation factor Tu	ND	ND	2
RPA3253	fusA, EF-G	elongation factor G	62	63	2
RPA3618		AAA ATPase	ND	ND	1
RPA3834	icln	NADP-dependent isocitrate dehydrogenase	20	43	2
RPA3876	fumA	fumarate hydratase, class I	45	41	2
RPA3878		conserved unknown protein	2	25	1
RPA4048	rbsF	alpha-D-glucose-1-phosphate	8	30	1
RPA4049	rbsG	cdp-glucose 4,6-dehydratase	16	50	1
RPA4050		unknown protein	18	62	1
RPA4433	clpB	Clp endopeptidase ATP-binding subunit	73	61	2
RPA4464	sorC	molybdopterin subunit sulfite oxidase	15	36	1
RPA4465	sorB	sulfite dehydrogenase	17	42	1
RPA4618	nifK	nitrogenase beta subunit	19	47	1
RPA4619	nifD	nitrogenase alpha subunit	38	63	1
RPA4620	nifH	nitrogenase iron-protein	12	51	2
RPA4641	cbbM	RuBisCo Form8	34	74	4
RPA4644	cbbP	phosphoribulokinase	33	76	2

ND: not detected by mass spectrometry

Table 2: Isolation of Nitrogenase complex

- Three proteins in a nitrogenase complex (nifD, nifH, nifK) were each cloned with both His6 and V5 epitope tags, expressed under photoheterotrophic, nitrogen-fixing conditions, and isolated.
- LC-MS-MS analysis of the isolate from each labeled component showed evidence for all three components of the nitrogenase complex.

Tagged Protein:	nifK seq cov. (%)	nifK unique peptides	nifD seq cov. (%)	nifD unique peptides	nifH seq cov. (%)	nifH unique peptides
nifK	**47**	**19**	**54**	**31**	**6**	**2**
nifD	**41**	**19**	**63**	**38**	**14**	**4**
nifH	**22**	**4**	**17**	**3**	**29**	**9**

Table 3: Isolation of RNA Polymerase complex

- The alpha subunit of a DNA-directed RNA polymerase complex was cloned with both His6 and V5 epitope tags, expressed in *R. palustris*, and isolated.
- LC-MS-MS analysis of the isolate showed evidence for the beta, beta', and omega subunits of the complex. Additionally, two sigma factors for this polymerase were observed. These sigma factors are sub-stoichiometric, and associate with the polymerase under different conditions of growth or stress.

Gene	Description	unique peptides	sequence coverage (%)
RPA3226 (tagged)	DNA-directed RNA polymerase alpha subunit	37	73
RPA2692	RNA polyerase omega subunit	4	53
RPA3268	RNA polymerase beta subunit	50	42
RPA3267	RNA polyerase beta' subunit	28	26
RPA1288	RNA polymerase sigma subunit	3	11
RPA0367	Stress response sigma factor	2	6

CONCLUSIONS

- **Significant progress is reported in expression, isolation, and detection of protein complexes in *R. palustris*.**
- **Development of more robust methods for identifying proteins associated with the tagged protein is underway, based on the expanding, well-characterized background library that we are amassing.**
- **Future plans include:**
  - **Implementation of improved separations techniques, online with mass spectrometry, that provide wider dynamic range and potentially faster throughput**
  - **Integration with high-mass-accuracy measurements of both peptides and intact proteins.**
  - **Development of digestion protocols that are faster, amenable to smaller sample sizes, and more compatible with subsequent MS analysis. [10]**

REFERENCES

1. Larimer, F. W. *et al.*, "Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*." *Nature Biotech.* **2004**, *22*, 55-61.
2. http://www.genomestolife.org
3. Buchanan, M. V. *et al.*, "Genomes to Life 'Center for Molecular and Cellular Systems:' research program for identification and characterization of protein complexes." *OMICS* **2002**, *6*, 287-303.
4. (a) Gavin, A.-C. *et al.*, "Functional organization of the yeast proteome by systematic analysis of protein complexes," *Nature* **2002**, *415*, 141-147.
   (b) Ho, Y. *et al.*, "Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry," *Nature* **2002**, *415*, 180-183.
5. Verberkmoes, N.C. et al., "Integrating "top-down" and "bottom-up" mass spectrometric approaches for proteomic analysis of *Shewanella oneidensis*," *J. Proteome Res.* **2002**, *1*, 239-252.
6. Strader, M.B., *et al.*, "Analysis of the 70S ribosome from *Rhodopseudomonas palustris* using integrated top-down and bottom-up mass spectrometric approaches." Sixth International Symposium on Mass Spectrometry in the Health and Life Sciences: Molecular and Cellular Proteomics, San Francisco, CA, Aug. 24-28, 2003.
7. Puig, O. *et al.*, "The tandem affinity purification (TAP) method: a general procedure of protein complex purification," *Methods* **2001**, *24*, 218-229.
8. Link, A.J. *et al.*, "Direct analysis of protein complexes using mass spectrometry," *Nature Biotech.* **1999**, *17*, 676-682.
9. Eng, J. K. *et al.*, "An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database." *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976-89.
10. Russell, W. K. *et al.*, "Proteolysis in Mixed Organic-Aqueous Solvent Systems: Applications for Peptide Mass Mapping Using Mass Spectrometry." *Anal. Chem.* **2001**, *73*, 2682-85.

ACKNOWLEDGEMENTS

This research sponsored by the Office of Biological and Environmental Research, U.S. Department of Energy. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.